



Quinn, N., Bates, P., Neal, J., Smith, A., Wing, O., Sampson, C., Smith, J., & Heffernan, J. (2019). The spatial dependence of flood hazard and risk in the United States. *Water Resources Research*, 55(3), 1890-1911. <https://doi.org/10.1029/2018WR024205>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1029/2018WR024205](https://doi.org/10.1029/2018WR024205)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via AGU at <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018WR024205> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Water Resources Research



RESEARCH ARTICLE

10.1029/2018WR024205

Key Points:

- 1,000 years of realistic U.S. flood patterns, comprising >63,000 individual events, are simulated using a statistical model
- Monetary losses for each event are calculated using a continental hydrodynamic model at 30 m resolution
- The analysis suggests that there is a 1% chance of U.S. annual fluvial flood losses exceeding \$78Bn in any given year

Supporting Information:

- Supporting Information S1

Correspondence to:

P. D. Bates,
paul.bates@bristol.ac.uk

Citation:

Quinn, N., Bates, P. D., Neal, J., Smith, A., Wing, O., Sampson, C., et al. (2019). The spatial dependence of flood hazard and risk in the United States. *Water Resources Research*, 55. <https://doi.org/10.1029/2018WR024205>

Received 5 OCT 2018

Accepted 3 FEB 2019

Accepted article online 13 FEB 2019

The Spatial Dependence of Flood Hazard and Risk in the United States

Niall Quinn¹, Paul D. Bates^{1,2} , Jeff Neal^{1,2} , Andy Smith¹, Oliver Wing^{1,2} , Chris Sampson¹, James Smith³ , and Janet Heffernan⁴

¹Fathom Ltd., The Engine Shed, Station Approach, Bristol, UK, ²School of Geographical Sciences, University of Bristol, Bristol, UK, ³Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA,

⁴J. Heffernan Consulting, UK

Abstract In this paper we seek to understand the nature of flood spatial dependence over the conterminous United States. We extend an existing conditional multivariate statistical model to enable its application to this large and heterogeneous region and apply it to a 40-year data set of ~2,400 U.S. Geological Survey gauge series records to simulate 1,000 years of U.S. flooding comprising more than 63,000 individual events with realistic spatial dependence. A continental-scale hydrodynamic model at 30 m resolution is then used to calculate the economic loss arising from each of these events. From this we are able to compute the probability that different values of U.S. annual total economic loss due to flooding are exceeded (i.e., a loss-exceedance curve). Comparing these data to an observed flood loss-exceedance curve for the period 1988–2017 shows a reasonable match for annual losses with probability below 10% (e.g., >1 in 10-year return period). This analysis suggests that there is a 1% chance of U.S. annual fluvial flood losses exceeding \$78Bn in any given year, and a 0.1% chance of them exceeding \$136Bn. Analysis of the set of stochastic events and losses yields new insights into the nature of flooding and flood risk in the United States. In particular, we confirm the strong relationship between flood affected area and event peak magnitude, but show considerable variability in this relationship between adjacent U.S. regions. The analysis provides a significant advance over previous national flood risk analyses as it gives the full loss-exceedance curve instead of simply the average annual loss.

Plain Language Summary Traditional flood risk analyses make the assumption that flow probability (the chance that a given river discharge is exceeded) does not vary within river catchments within an event. Real floods, however, do not look like this: In some places flooding is more severe than in others. Over a few tens of kilometers of river assuming the same event return period everywhere is perfectly fine, but over larger areas it breaks down. At national scales traditional risk analyses can only estimate the average annual loss. To estimate the total annual losses that might occur in more extreme flooding years the risk analysis needs to be based on more realistic spatial patterns of flooding. In this paper we use a sophisticated statistical model, based on U.S. Geological Survey river flow data, to simulate 1,000 years of spatially realistic U.S. flooding comprising more than 63,000 individual events. By calculating the damage for each event as a dollar value, we are able to estimate the probability of the United States experiencing particular levels of annual flood losses. We show that there is a 1% chance of U.S. annual fluvial flood losses exceeding \$78Bn in any given year, and a 0.1% chance of them exceeding \$136Bn.

1. Introduction

While intense and isolated thunderstorms can cause extreme flows and flooding that affect only single river gauging sites (e.g., the Hepner storm of 1903 in Eastern Oregon and the Boscastle flood of 2004 in the United Kingdom), extreme rainfall more typically occurs as a result of larger-scale and organized weather systems which can cause flooding in a number of different basins along the storm track. Numerous river gauging sites in a region may therefore experience extreme flows as a result of the same weather event. Because rainfall patterns and hydrological processes have distinct spatial patterns, the extreme river flows that these processes generate will also have a spatial structure. In particular, the spatial pattern of maximum flood return periods caused by a single weather event at a set of gauge sites across a region is usually characterized by spatial dependence: If we observe an extreme flow at a particular site there is a positive probability that nearby sites will also be experiencing high flows as a result of the same weather event (i.e., they are

©2019. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

spatially auto-correlated). “Nearby” in this case could either mean within the same river basin or in proximal basins, and in addition implies that there will be some distance decay in the spatial dependence between sites as the separation between them increases. This dependence structure is therefore asymptotic (Keef et al., 2012), as in the limit where the separation distance between sites approaches zero, the probability that two sites experience the same flood return period approaches 1. In physical terms this dependence structure is a result of the interactions between the space-time structure of rainfall (Singh, 1997; J. A. Smith et al., 1994), land surface properties that affect runoff generation (Blöschl & Sivapalan, 1997), network topology (Gupta et al., 1996), and channel-floodplain hydraulics (Turner-Gillespie et al., 2003). Spatial patterns of flooding are also characterized by asymptotic independence, whereby for very distant sites there will be zero probability that the highest flows will be caused by the same storm (Keef et al., 2012). This is because there are physical limits to the maximum size of storm systems, although this may be very large for major regional flooding events (e.g., Mississippi 1993, Central Europe 2003 and 2013, Pakistan 2010, Bangladesh 2011). The spatial pattern in flood return period for an event is commonly referred to as the event footprint (e.g., Rougier, 2013).

Despite these spatial properties of floods being relatively well known (at least in general terms), they are substantially understudied compared to the at-a-point case. In the United Kingdom, Keef et al. (2009), in a study looking at both extreme precipitation and river flows, found that spatial dependence was stronger for the latter and hypothesized that this was because precipitation was more influenced by small-scale variability, for example, as a result of atmospheric convection. Keef et al. (2009) also found that areas of the United Kingdom with diverse catchments exhibited lower levels of flood peak spatial dependence and that this catchment variability dominated over the dependence pattern caused by large-scale precipitation gradients. Spatial dependence patterns also became less smooth with increasing event return period (Keef et al., 2009).

In the United States, Villarini et al. (2011) demonstrated asymptotic dependence and strong spatial variability for flood peaks for more than 2,000 U.S. Geological Survey (USGS) river discharge gauges during hurricanes Frances, Ivan, and Jeanne from 2004 and hurricane Hanna in 2008. As a result, they recommended that regional-scale flood models needed to move away from at-a-point flood frequency analysis and should incorporate the spatial correlation of flood peaks. Villarini and Smith (2010) showed strong spatial variability in the parameters of the Generalized Extreme Value (GEV) distribution for 572 USGS gauging stations on the U.S. eastern seaboard, implying the existence of significant spatial dependence within flood events. They also found striking spatial heterogeneity in the frequency of flood peaks caused by tropical cyclones, with maxima in south Florida and along the Atlantic seaboard between north Georgia and New Jersey and a local minimum in between. The existence of different flood generating mechanisms in different U.S. regions (e.g., J. A. Smith et al., 2011) also strongly suggests that spatial variation in dependence is likely to be a ubiquitous phenomenon. J. A. Smith et al. (2011) found clear spatial heterogeneities in rainfall and flood peak distributions in the eastern U.S. seaboard region due to a variety of factors including orographic precipitation enhancement, local atmospheric circulations near land-ocean boundaries, and urbanization impacts on regional climate. Lastly, Lu et al. (2017) examined spatial variability in flood magnitude for four floods events in the Delaware river basin in the north-eastern United States, simulated using a distributed hydrologic model that was forced by high-resolution radar rainfall fields. They found asymptotic dependence, with decreasing return period correlation with increasing separation distance, but with strong variation between flood events. Despite the four events having similar peak discharge at the basin outlet, they found that the distributions of flood magnitudes over the drainage network varied markedly. For the small sample of events studied they found that more extreme flows were linked to stronger downstream correlation along the main stem, which was hypothesized to be the result of (i) flood wave attenuation by channels and floodplains and (ii) aggregation of flows at tributary junctions.

These studies begin to uncover the spatial dependence in peak flows during flood events, but further work is required to systematically document spatial dependence and how this relates to geology, meteorology, catchment structure, and vegetation patterns, particularly over large and heterogeneous regions such as the United States. Previous work in the United States has only looked at a limited number of specific storms (Villarini et al., 2011) or over a few catchments (e.g., Cooley et al., 2007; Lu et al., 2017), and there is thus considerable scope for a more comprehensive analysis.

This limited number of studies which address the spatial dependence in flood peaks at multiple locations during events contrasts markedly with the substantial body of work which considers at-a-point (univariate) flood frequency analysis. The latter is a standard technique in engineering hydrology and is routinely employed to manage flood risk at individual sites. Univariate extreme value theory is used in both gauged and ungauged catchments to estimate return period flows for such tasks as designing flood defenses, flood-plain mapping, and risk analysis. These tasks are typically conducted over relatively small spatial scales (a few tens of kilometers of river reach) such that the assumption that the flow return period is constant in space is an acceptable one. However, there are other classes of flood risk management decision where this “constant in space” return period assumption breaks down, and the effect of spatial dependence on flood frequency needs to be taken into account. Prime examples are to be found in the insurance and reinsurance industries. Individual insurance firms need to know the probability distributions of annual losses to which their portfolio is exposed (i.e., the so-called loss-exceedance curve) in order to reassure market regulators that they can maintain solvency in the face of rare sequences of flood events. Because insurance firms operate regionally and nationally, the risk analysis is required at this scale. Typically, regulators require evidence that firms can absorb the 0.5% annual probability (1 in 200-year return period) loss and still continue trading. Reinsurance is effectively insurance for insurers to help them withstand high loss years and has similar requirements for information. In both cases treating the flood frequency at a point as if it were independent of all other points will misestimate the potential losses (Timonina et al., 2015). Univariate extreme value analysis conducted at a series of sites independently cannot generate the year-to-year variation in flood losses that insurers experience (e.g., Pielke et al., 2008), and a multivariate analysis considering spatial dependence is required. Governments may also need to know the probability of particular annual national losses due to flooding in order to develop the economic case for investment in flood defenses, and emergency services need to know how likely and widespread regional-scale flooding events will be to ensure sufficient resources can be mobilized in good time and at the right locations. Lastly, civil contingency planners need to know about the likelihood of plausible events that have not yet been seen in the short instrumental record we have for river flow. It is important to think about such “black swan” events (Taleb, 2007) in preparing national risk registers such as the one compiled by the U.K. Government’s Cabinet Office (Cabinet Office, 2017).

Almost all current national-scale flood risk mapping by governments (e.g., FEMA’s Special Flood Hazard Area mapping program in the United States or the U.K. Environment Agency’s Flood Map program) creates a national return period flood hazard map from a patchwork of local modeling studies, each of which uses the constant in space return period assumption. However, when this patchwork is assembled into a single national map this assumption breaks down (e.g., Jongman et al., 2014). By intersecting a large sample of different constant in space return period hazard maps with information on asset locations and their vulnerability, the potential flood exposure and national-scale expected Average Annual Loss (AAL) can be calculated (e.g., Wing et al., 2018), but it is not possible to calculate an accurate risk profile (i.e., the loss-exceedance curve) because of the lack of spatial dependence. Figure 1 illustrates this idea graphically, and a flow chart describing each calculation chain is given in the supporting information (Figure S1).

The same limitation also holds for large-scale flood inundation models which use univariate flood frequency analysis to define T year return period flows along river networks (Alfieri et al., 2013; Dottori et al., 2016; Pappenberger et al., 2012; Sampson et al., 2015; UNISDR, 2015; Winsemius et al., 2013; Yamazaki et al., 2011). Large-scale modeling studies forced by observed gauged flows (Schumann et al., 2016) can capture elements of the historic spatial dependence (to the extent the forcing data represents this). However, the 40- to 50-year length of typical gauged records is not sufficient to simulate all possible patterns of flooding (Serinaldi & Kilsby, 2017). As a result, most large-scale flood risk analyses published to date do not account for spatial dependence and hence cannot address the national-scale flood risk management questions posed above. A few recent academic studies, for example, by Falter et al. (2015), Timonina et al. (2015), and Jongman et al. (2014), have highlighted the need to account for interdependencies in flooding between regions when estimating risk and present approaches to do so. However, assessments used by governments do not yet adequately represent spatial dependence. Anecdotally at least, it is known that flood risk models used in the insurance industry do include spatial dependence. However, these approaches are commercially privileged and exactly how this is done is not clear as these methods are rarely subject to open and independent scrutiny via peer-review.

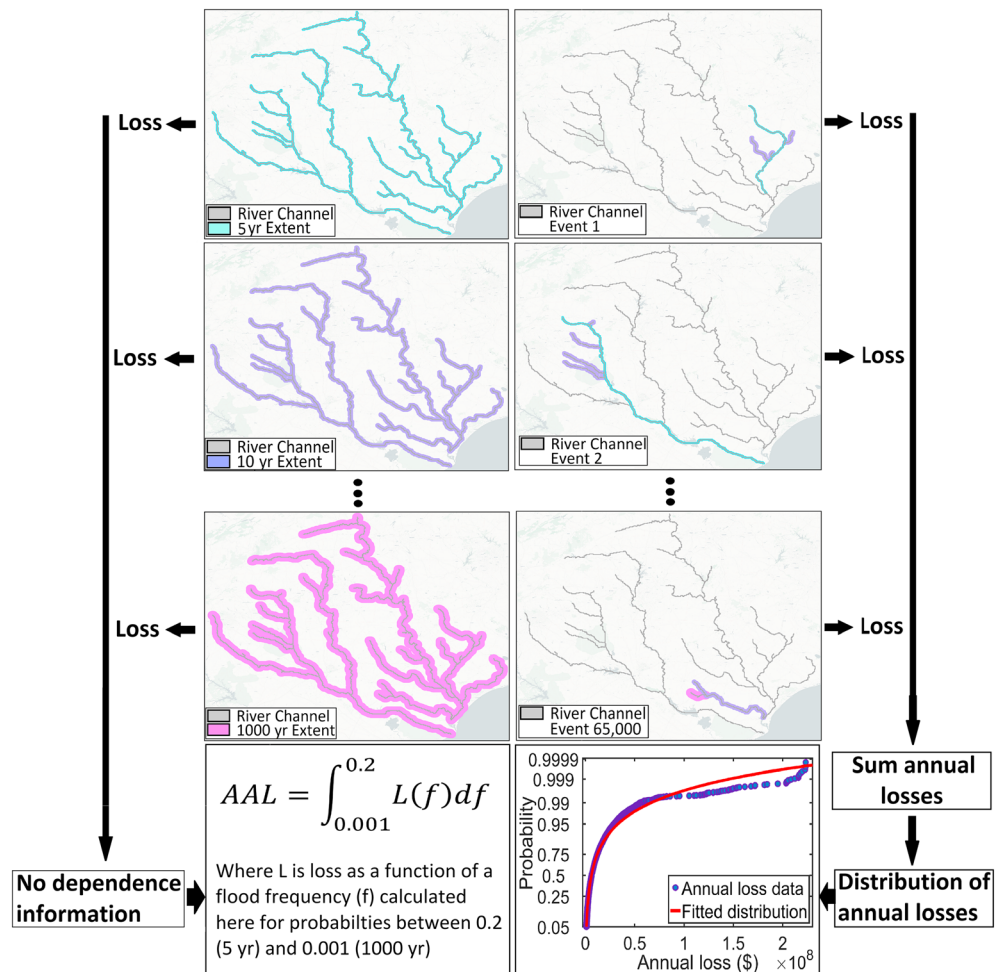


Figure 1. A schematic visualizing the difference between the traditional large-scale flood risk estimation method and that proposed in this research. The left panel (the traditional approach) shows a series of static hazard layers for a region (a single basin in this case), each of which are intersected against asset data and depth-damage functions to estimate losses for each return period. The given return period flow is assumed to occur everywhere in the region simultaneously. From this only AAL can be calculated, and this is then given as the integral of the loss-probability curve. As no information is known relating to likely events and their spatial pattern it is impossible to know the range of losses one might experience over a given time period. The right panel (the proposed method) simulates many thousands of individual synthetic flood events that have realistic spatial patterns. For each event, losses are estimated in the same way as in the traditional method. Event losses are then assigned to the years in which they occur to provide a distribution of annual (sum) losses. This distribution reveals the full annual loss-probability curve and, therefore, the likelihood of any given year exceeding a particular loss threshold. AAL = Average Annual Loss.

The methods available to investigate the spatial dependence in peak flows during flood events break down into (i) empirical analyses of spatial dependence in gauged records (e.g., Villarini et al., 2011) or model output (e.g., Falter et al., 2015; Lu et al., 2017) or (ii) studies which attempt to fit statistical models of the dependence to large sets of gauge records containing multiple events (Heffernan & Tawn, 2004; Keef et al., 2009). The latter approach provides a more comprehensive view of the dependence structure and also has the advantage that it allows the stochastic simulation of new event footprints from the statistical model, including plausible events not previously seen in the gauged record. In this way, tens of thousands of likely events can be generated and fed into hydraulic models to produce hazard maps that correctly incorporate spatial dependence. In turn, this then allows better sampling of the tails of the risk distribution and overcomes the problems with current large-scale flood mapping approaches noted above.

Broadly, two classes of statistical method can be used to study spatial dependence in environmental data: copulas (e.g., Chen Lu et al., 2012; Nelsen, 2006) and multivariate (many site) extreme frequency analysis

(e.g., Heffernan & Tawn, 2004). Both methods compute spatial dependence, however copulas do not allow the dependence to change as flows become more extreme (Keef et al., 2012). Previous authors (Keef et al., 2009) have noted that extreme flows often show weaker spatial dependence than normal river flows as the latter may be more influenced by large-scale, long-term climatic patterns rather than localized weather systems. Extreme flows also show both asymptotic dependence and asymptotic independence (Keef et al., 2009) which is explicitly represented within models such as the one developed by Heffernan and Tawn (2004). As a result, multivariate extreme models may be a better choice of statistical model for understanding spatial dependence during flood events.

The aim of this paper is therefore to undertake a comprehensive investigation of spatial dependence during floods events within a large and heterogeneous region using conditional multivariate extreme value analysis. The spatial dependence structures generated are then used to produce many plausible synthetic event footprints, which, when intersected with asset and depth-damage information, are used to predict economic loss (and risk) to assets across the conterminous United States. We select the conterminous United States for this study because of the quality of gauging station and other data, and also because of the diversity of climates, landscapes, and thus flood generating mechanisms that exist (e.g., O'Connor & Costa, 2004; J. A. Smith et al., 2011; Vogel et al., 2001). Previous applications of multivariate extreme value techniques to flooding have only considered relatively small and homogeneous regions (e.g., the United Kingdom and Austria: Keef et al., 2009, 2012; Schneeberger et al., 2018; Speight et al., 2017; Wyncoll & Gouldby, 2013), and as a result have only been applied to sets of a few hundred gauging sites at most. An analysis for the United States will need to consider many thousands of gauge sites and this gives a significantly more complex problem for the analysis. First, as we need to compute correlations between each gauge and the m other sites in the network the task scales with m^2 giving a large compute problem. Second, over such a large region the problem of deciding what constitutes a unique event requires careful handling. In section 2 below we outline the conditional multivariate extreme value analysis used and the methodological developments required to address the challenge of applying such methods to the United States. In section 3 we present results, and in section 4 we discuss what these tell us about the nature of flood spatial dependence and risk in the United States. Conclusions are provided in section 5.

2. Method and Application to the Continental United States

In flood risk studies, it is common to characterize flow magnitudes by their “return period,” usually with respect to years. This term defines the length of time, on average, between flow magnitudes exceeding some threshold based on the annual flow exceedance probability (the likelihood a given flow is exceeded in any given year). For instance, a flow that exhibits an annual exceedance probability of 0.1 (10%) would be expected to have a return period of 1 in 10 years. Due to its common use, particularly among end users of flood risk analysis studies, this study uses the return period description when discussing extreme flows.

For a flow exceeding a specified threshold, T , at gauging site X , $Q_X > T$, where T is large enough to cause flooding, we seek to determine the probability distribution of flow at gauge Y_i , that is, $\Pr(Q_{Y_i}|Q_X > T)$ for all m gauges in a set where $i = \{1, 2, \dots, m\}$. Following Keef et al. (2009) we use the conditional exceedance model of Heffernan and Tawn (2004) to perform this calculation. This approach has proved robust and flexible in previous applications to flooding (Keef et al., 2009, 2012; Lamb et al., 2010; Wyncoll & Gouldby, 2013) as it enables the prediction of events beyond the range of the data used to train it and does not assume variables are only asymptotically dependent. The statistical model therefore provides the distribution of a set of variables (river flow in this case) conditioned upon one of the members of the set being extreme, where “extreme” is a user defined threshold.

Application of this approach to river gauging sites in the United States requires a number of extensions to previous applications of the method. First, developments are required to deal with the significantly increased computational burden of application to such a large and heterogeneous domain. The previous largest number of gauges analyzed was ~300 for an application in the (relatively homogeneous) United Kingdom by Keef et al. (2009), yet for the United States a set of many thousands of gauges will be required to characterize spatial dependence. Given that dependence is a pairwise calculation, the computational problem scales with m^2 where m is the number of gauges. We solve this problem by parallelizing the calculations of dependence for each conditioning site over a large computer cluster of >800 cores, that is, every gauge is considered

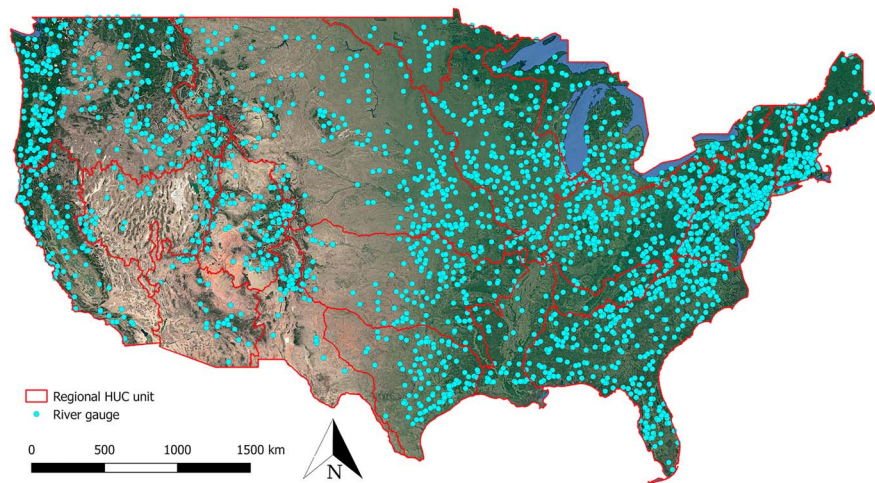


Figure 2. Location of river gauge records used in the analysis that passed the quality check procedures described in the text. The underlying HUC units used to disaggregate the United States into 18 large regions are also shown.

independently, calculating the dependence structure to other gauges. Second, most published applications of the Heffernan and Tawn (2004) method have simply computed the dependence structure at a set of gauges (Keef et al., 2009, 2012; Lamb et al., 2010) and then used this to stochastically simulate new events at the gauge sites (an “event set”). However, to calculate flood risk we additionally need a method to translate the simulated flows at gauging sites into regional maps of flood extent and depth so that the loss can be computed. Following Wyncoll and Gouldby (2013) this is achieved using a library of simulations produced by a numerical flood inundation model. However, instead of the simple flood spreading method applied to a single catchment demonstrated by Wyncoll and Gouldby (2013), we here use a two-dimensional hydrodynamic model of the entire continental United States at 30 m resolution (Sampson et al., 2015; Wing et al., 2017). Standard depth-damage functions developed by the U.S. Army Corps of Engineers can then be used to estimate the monetary loss that results from each flood footprint. Use of a precomputed simulation library also significantly increases the efficiency of the method compared to running each separate event through the hydraulic model, which would be highly expensive even on the largest supercomputing cluster. Lastly, given the size of the region it is likely that a flood wave may take several days to pass between neighboring gauges. Keef et al. (2009) outlined a lag-based dependence approach based on the work of Heffernan and Tawn (2004) that has been used in numerous studies since. We use the same approach here but adapt it for large domain areas using the USGS Watershed Boundary Data Set (<https://nhd.usgs.gov/wbd.html>) level 2 HUC basins to define 18 large regions across the United States that are hydrologically connected (see Figure 2). We then use this information to constrain the number of gauges that can experience the same event. This helps to reduce the occurrence of spurious correlations when calculating dependence and reduces significantly the number of pairwise dependence calculations that are required. We use this as opposed to a simple distance measure as this would not take hydrological connectivity into account.

2.1. Gauge Data

Daily river flow time series were obtained from the USGS river gauge network (<https://waterdata.usgs.gov/nwis/>). This data set contains approximately 9,000 gauge records of varying length and quality. To ensure consistency we select only sites with a 40-year record from 1976 to 2016 that contain fewer than 2% erroneous or missing data and which have no obvious step change or trend in extreme flows over the measurement period, as determined by a Kendall Tau rank correlation coefficient test. Data gaps within the set of retained sites were infilled using regression relationships with neighboring gauge time series, as current multivariate modeling approaches such as the Heffernan and Tawn (2004) model are unable to handle missing data. The resulting data set comprises 2,400 high quality, long-term time series at the locations shown in Figure 2. The selection thresholds may be viewed as somewhat conservative, however they were used to reduce the number of missing flow values that required infilling in order to minimize the likelihood of introducing spurious flow estimates into the data used to create the spatial dependence model.

As basin size decreases the rainfall-runoff response becomes increasingly flashy and the mean daily flow will become an increasingly imperfect estimate of the maximum daily discharge. However, sufficient subdaily temporal resolution flow data do not exist in the United States with which to produce a national-scale dependence model. To improve computational efficiency when dealing with temporal dependence in flows at different sites, the daily data were converted to time series of 3-day maxima. However, as we compute dependence over a 9-day window the effect of this on the results is likely to be small. Gauge records along the lower Mississippi (extending from Louisiana to Kentucky) were sparse within the USGS discharge data set. For this reason, a selection of supplementary stage records were extracted from an independent data set provided by the U.S. Army Corps of Engineers. Stage data sets could be treated in exactly the same way as discharge data within the analysis as the method uses the marginals of each distribution rather than the actual values. Assessment of sites where both stage and discharge measurements existed indicated a very high level of correlation between these two variables, indicating that the use of either variable in the statistical model is acceptable.

2.2. Statistical Model

In the method of Heffernan and Tawn (2004) the conditional exceedance calculation comprises two independent steps: First, the marginal probability distribution is defined for each gauge site, then the dependence structures between sites are computed as a series of pairwise regressions. Full details can be found in Heffernan and Tawn (2004) and Keef et al. (2009, 2012).

The marginal distributions were calculated independently at each gauge site. They define the probability a gauge will exceed a given flow magnitude and are calculated using a semi-parametric function that fits a generalized Pareto distribution to discharge values above a specified quantile threshold and an empirical distribution to those below it. The specified threshold should be large enough so that the spatial dependence assumption holds, but low enough that there are sufficient data above the threshold with which to estimate the generalized Pareto model (Neal, Keef, et al., 2012). A number of tests were performed to define the marginal fitting threshold. First, QQ plots were used to define errors between the fitted generalized Pareto distribution and the observed data under different threshold specifications. Next, the stability of the generalized Pareto scale and shape parameters given a change in threshold specification was assessed. Following this approach, the 92nd quantile of the 3-day maximum time series was found to routinely provide the lowest errors within the QQ plots while also providing relatively stable Pareto parameter estimates across the gauge sites sampled.

As this study is interested in defining the dependence between extreme flows within given events, some spatial and temporal limitations need to be set to define what we consider to be an independent event. This is a complex task given the heterogeneity of flood hydrology across the United States (J. A. Smith et al., 2011) and some simplifications are necessary based on broad spatial and temporal limits. First, a spatial limit is put upon event size using 18 large hydrological regions (the level 2 USGS HUC basins, see Figure 2). We allow dependence to be calculated from a conditioning site to any other gauges that are within its own region or any region adjacent to it. All other gauges are assumed to have zero dependence to the conditioning site in question. In addition, a temporal window of 9 days is implemented to account for the variability in arrival time of a flood peak between sites experiencing the same event. This window was selected as even apparently long duration U.S. floods, such as the 1993 Mississippi event, consist of multiple flood waves each of which would comprise a distinct event in our framework. Moreover, a comprehensive analysis of flood wave travel times by Allen et al. (2018) shows that, globally, flood waves take a median travel time of 6, 3, and 2 days to reach their basin terminus, the next downstream city, or the next downstream dam, respectively. Thus, the time taken for a flood wave to cross any adjacent pair of level 2 HUC basin units is likely to be <9 days. While the choice of spatial and temporal limits here is necessarily somewhat arbitrary, the above decisions are sensible given current knowledge of the typical duration and spatial dependence of U.S. floods. The modeling approach is highly flexible and future applications could implement different spatial and temporal windows should evidence emerge that these are required.

Once the marginal distributions and time/space limits are specified we can then calculate the dependence between gauges. The first step is to transform the marginal data onto a common scale. The Laplace distribution was used in this study after Keef et al. (2009) who highlight that it provides a more appropriate linear

relationship between extreme observations than the Gumbel distribution. The dependence model can be described by equation (1) (after Wyncoll & Gouldby, 2013):

$$\mathbf{Y}_{-i}|Y_i = \mathbf{a}Y_i + Y_i^b \mathbf{Z}_{-i} \quad \text{for } Y_i > v, \quad (1)$$

where \mathbf{Y}_{-i} is a vector of the marginal distributions at each gauge excluding gauge Y_i , v is a high threshold above which dependence is modeled, \mathbf{a} and \mathbf{b} are vectors of parameters (limited to $-1 < \mathbf{a} < 1$ and $\mathbf{b} < 1$, respectively) describing the strength of the dependence and how it changes with increasing magnitude of Y_i , and \mathbf{Z}_{-i} is a vector of residuals. Equation (1) is implemented in a pairwise manner so that the j th element of \mathbf{Y}_{-i} is modeled as a function of Y_i using parameters $a_{j|i}$, $b_{j|i}$, and residuals $Z_{j|i}$, while the dependence between components of \mathbf{Y}_{-i} are modeled nonparametrically using the joint distribution of the residual $Z_{j|i}$.

As noted previously, flood peaks at different locations during an event typically do not occur concurrently. Therefore, the temporal lag between gauges is represented using the approach defined by Keef et al. (2009) in which dependence is calculated between gauges at all temporal lags within the specified time window. Therefore, the conditional model becomes

$$Y_{j,t+\tau}|Y_{i,t}, \quad (2)$$

where τ is a selection of temporal lags.

The drawback to this approach is that it can result in a significant increase in computational requirements and is a further reason why in this study the daily time series are coarsened to 3-day maxima. This means that dependence only needs to be calculated at two additional “lag time series” for each gauge in the network over the 9-day time window.

2.3. Simulation of an Event Set of Spatially Dependent Gauge Flows

The conditional multivariate model characterizes the dependence between sites when the flow at one or more of the sites in the set of gauges is extreme. This information can then be used to generate a synthetic catalogue of events that samples the tails of the flood risk distribution better than the limited observational record. To do this, the observational record was used to define the number of events expected to occur across a region in any given year. The 99th quantile of the flow distribution was used to define an extreme event at any given site, while the temporal and spatial windows (discussed above) were used to group these flows into independent events. An empirical distribution was then fitted to the annual event counts. For each year of simulation, the number of events to be generated was sampled from this distribution.

To establish where to sample for a given simulated event (i.e., which gauge site should be conditioned upon) we first calculate the relative likelihood that any gauge is the largest (on the Laplace margin) using the observational record. Once the likelihood that any gauge is the largest during a given event is estimated, multinomial sampling is then used to select the gauge to condition upon. For each conditioning site the dependence model is used to simulate return periods at all gauged locations. The magnitude of the event at the conditioning site is sampled above a specified threshold, set to the 99th quantile in this study. To ensure self-consistency within the model, a rejection sampling scheme was used in which events are generated at the conditioning site until one is produced in which the conditioning site is the largest out of all locations (in terms of the Laplace margins). Within this study this process was repeated for 1,000 years, resulting in an ensemble of approximately 63,000 flood events where at least one gauge has simulated flow greater than the 99th quantile. Approximately 42,000 of the events in the ensemble have at least one gauge with flow greater than 1 in 5-year return period. An example of the observed and simulated dependence between three gauges within the Platte River catchment in Nebraska is given in Figure 3. Here the dependence between a downstream gauge site (Gauge 1) and two further gauges, one proximal (Gauge 2) and one distal (Gauge 3), is calculated. As might be expected, the statistical model shows a strong dependence relationship between Gauges 1 and 2, but a much weaker relationship between Gauges 1 and 3. Figure 3 indicates that the statistical model can recreate the observed relationship above the specified prediction threshold in both instances despite the significant difference in dependence strength.

Given our spatial and temporal assumptions used to define “events,” the observed record contains approximately 2,500 distinct floods within the 40 years of data considered. By using stochastic sampling we are

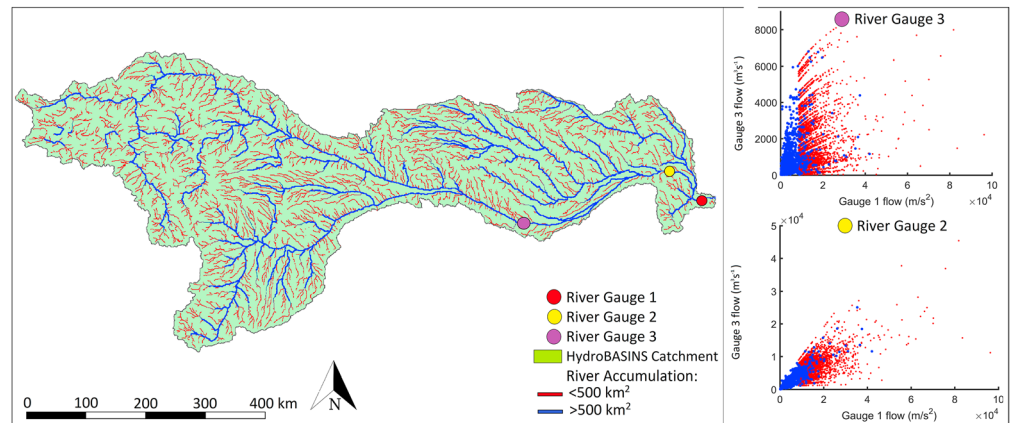


Figure 3. An example of tail dependence between river gauges along the Platte river in Nebraska, west of Omaha (HydroBASINS level 5 catchment). Scatterplots of modeled flows (red) and observed flows (blue) at two neighboring gauges (Gauges 2 and 3) relative to flows at a conditioning gauge.

therefore able to dramatically increase the number of event data sets we have for analysis. There is no upper limit to the size of the event set that can be generated, and here we create a 1,000-year synthetic flood series with the same marginal distributions and dependence structure as the historic record. Using this allows us to better sample the tails of the distribution and creates the possibility of generating plausible extreme events that have not yet been seen within the limited gauge observations we have. The synthetic record will thus contain many examples of typical hydrological design floods (event peak flow magnitudes with 1% or 0.5% annual probability), along with sufficient numbers of even rarer events with which to characterize the distribution tail beyond this.

2.4. Flood Extent and Depth Maps

Each member of the event catalogue generated using the conditional multivariate model consists of estimates of flow return periods at a set of gauge points which then need to be converted into surfaces of flood extent and depth (the flood footprint) in order to perform a loss calculation. To do this for each event, we first define catchment areas around each gauge which experience the same extreme return period flow. We then determine the flood extent and depth corresponding to this return period within that zone. Interpolation of gauged flows to catchment areas is performed by first discretizing the United States into independent units defined by the HydroBASINS data set (Lehner & Grill, 2013). HydroBASINS is a series of hydrologically connected regions that are subdivided into increasingly smaller areas through a series of levels from 1 (largest) to 12 (smallest). To create flood extent and depth maps we use a hybrid layer based on the level 8 basins (see Figure 4), subdividing to level 10 basins where multiple gauges are found within a level 8 unit. An advantage of using the HydroBASINS data set to discretize regions is that there is a physical basis by which to expect units to experience the same event.

Hydrological units that contain river gauges are automatically assigned the return period calculated for that gauge. Ungauged units have their values inferred from the most appropriate neighboring gauge site value during a given event. This was done using a two-step process. First, catchment centroid locations and upstream accumulation areas were obtained for every catchment and gauging station. Second, using this information, a distance measure was calculated between each ungauged catchment and each gauge. The distance measure was given as the weighted sum of the normalized distance and accumulation errors between a catchment and each potential gauge. A Monte Carlo analysis was used to define the optimal weighting of the distance and accumulation errors using a leave-one-out approach, comparing modeled predictions with observed records. The most appropriate gauge from which to reference the expected event magnitude at an ungauged catchment was given as the one with the lowest overall weighted error. The use of a nearest neighbor-based approach was selected to avoid smoothing of event footprints, which can lead to a reduction in extreme events in ungauged catchments, particularly within regions where gauge density is low. However, a possible drawback with this approach is that it may result in larger than expected event footprints in gauge-poor regions, particularly during events that are highly localized in nature.

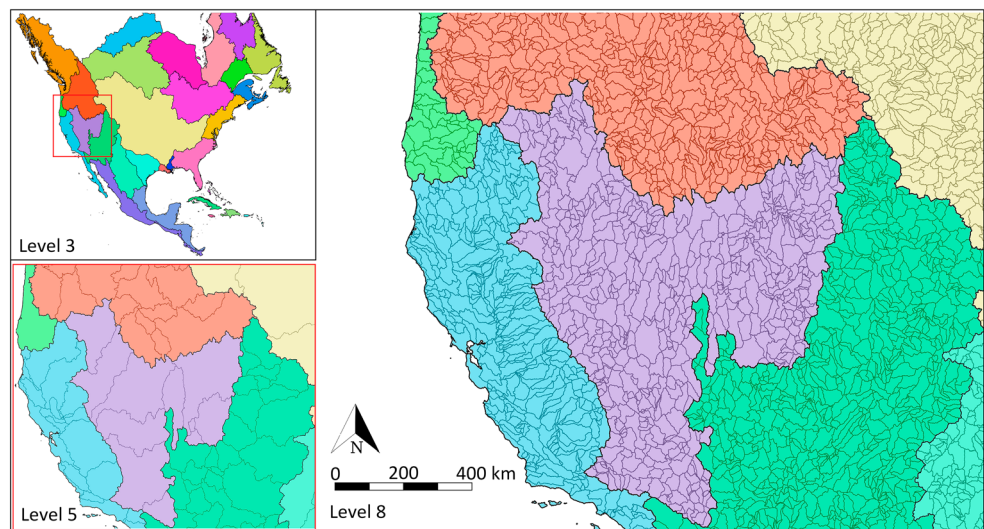


Figure 4. An example of the progressive discretization of the HydroBASINS catchment data set across three levels of complexity along the west coast of the United States.

Flood extent and depths associated with the event return period assigned to each hydrological unit were then extracted from a precomputed library of hydraulic simulations for the United States, generated using a two-dimensional hydrodynamic model (Sampson et al., 2015; Wing et al., 2017). This model is an implementation of the LISFLOOD-FP numerical scheme (Almeida et al., 2012; Almeida & Bates, 2013; Bates et al., 2010) deployed at 1 arc sec (~ 30 m) resolution over the entire continental United States. For this application river channels are delineated using the HydroSHEDS global hydrography data set (Lehner & Grill, 2013), while the floodplain is represented by a digital elevation model (DEM) derived from the 1 arc sec (~ 30 m) USGS National Elevation Data Set. Larger rivers are “burned” directly into the DEM, while smaller rivers are treated as subgrid scale features (Neal, Schumann, et al., 2012). The U.S. Army Corps of Engineers National Levee Database is incorporated into the model to explicitly represent known flood defenses and, because this is recognized to be incomplete, we utilize an additional defense approximation based on spatial variation in urbanization intensity defined in the National Land Cover Database of the Conterminous United States (Homer et al., 2015). Here we use four land cover categories relating to development level and assign protection standards to them, given as:

1. Developed, open space. <20% impervious surfaces. A 5-year protection standard is assigned.
2. Developed, low intensity. 20% to 49% impervious surfaces. A 20-year protection standard is assigned.
3. Developed, medium intensity. 50% to 79% impervious surfaces. A 50-year protection standard is assigned.
4. Developed, high intensity. $\geq 80\%$ impervious surfaces. A 100-year protection standard is assigned.

Therefore, within the most highly urbanized centers we assume that flood defenses (or equivalent resilience strategies) would protect against flooding up to the 1 in 100-year level, with decreasing levels of protection assumed as regions become less urbanized. Similarly, de Moel et al. (2015) and Jongman et al. (2014) also highlight a lack of comprehensive defense characteristics as a key issue when estimating flood risk. Friction parameters are uncalibrated and assigned based on land cover data (Homer et al., 2015) and sensible values from the literature (e.g., Chow, 1959).

The model simulates fluvial flooding in catchments down to 50 km^2 with boundary conditions derived from regionalized flood frequency analyses (RFFA; A. Smith et al., 2015). The RFFA uses approximately 6,000 USGS river gauges across the United States to determine design discharges for 17 different recurrence intervals (5-, 8-, 10-, 15-, 20-, 30-, 40-, 50-, 75-, 100-, 150-, 200-, 250-, 350-, 500-, 750-, and 1,000-year). A synthetic triangular hydrograph is then generated for each river reach based on a determined time to concentration. The fluvial component is only executed for catchments larger than 50 km^2 as there are insufficient stream-flow records to constrain the RFFA in areas smaller than this. For each catchment affected during an event we extract from the precomputed hazard library the flood extent and depth for the return period closest to the return period calculated by the above interpolation procedure.

Wing et al. (2017) validated the inundation extent predicted by the hydrodynamic model across the United States by comparing it to the entire catalogue of FEMA Special Flood Hazard Area maps and to the output from 10 hydraulic models built using high-quality local data by the USGS. Where the FEMA maps are based on high-quality local modeling, the continental-scale model attains an overall hit rate of 86%. The correspondence is better in temperate areas and in basins above 400 km² but is still reasonable for arid areas (hit rate ~73%) and smaller catchments (hit rate always >75%). While hit rate is not an ideal metric for evaluating flood models as it does not penalize overprediction, it is the only performance measure that can be used with the FEMA maps because of their patchy spatial coverage (see Wing et al., 2017, for a more complete explanation). Against the higher quality USGS data, the average hit rate reaches 92% for the 1 in 100-year flood, and 90% for all flood return periods and Wing et al. (2017) was also able to calculate the more informative Critical Success Index (or CSI) metric. Unlike the hit rate, CSI penalizes both underprediction and overprediction, and also ignores the bias caused by assessing the model performance over easy to predict nonflood-plain areas. Against the USGS data Wing et al. (2017) obtained CSI values from 0.51 to 0.90, with an average of 0.76 which suggests performance equal to that of typical reach scale models built using LiDAR terrain data by skilled operators and validated against satellite Synthetic Aperture Radar scenes (Bates & De Roo, 2000; Horritt & Bates, 2001a, 2001b). Given typical hydraulic modeling uncertainties in the FEMA maps and USGS model outputs (e.g., errors in estimating return period flows), it is probable that overall the continental-scale model can replicate both to within error.

2.5. Loss Calculations

Each event inundation and depth layer was overlain with asset information (location, characteristics, and value) for >111 million structures in the United States from the FEMA National Structure Inventory (http://data.fema.gov/data/FIMA/NSI_2010/). The monetary loss to these assets resulting from the simulated flooding was then calculated using depth-damage functions for different building types developed and utilized by Wing et al. (2018) based on the information provided by the U.S. Army Corps of Engineers. This data set is composed of a lookup table of loss relative to asset value for 40 unique asset types, across 29 flood depth bins. For instance, a residential property with one floor and no basement would be expected to incur a loss of 25% of its value at a flood depth of ~60 cm, increasing to 40% at ~1.8 m. The value of a given asset was taken from the National Structure Inventory data set. We assume that no loss occurs for predicted depths below 20 cm as this is a typical kerb height or property ground floor elevation above the land surface.

The output from this process is an expected economic loss within each catchment unit for each synthetic event footprint. An example flood footprint overlain onto the National Structure Inventory for an event in Washington state along the Skagit River is shown in Figure 5. For consistency all simulated and observed loss values reported in this paper are adjusted for inflation to 2018 values.

3. Validation of Simulated Event Footprints

To validate the stochastic event generator, we first visually inspected a random sample (~2%) of event footprints as a “sanity check” on the results. Figure 6 shows a subset of example footprints contained within the 1,000-year synthetic record, where points refer to catchment centroid locations. For clarity, catchments with <5-year return periods have been excluded while red dots represent those with ≥5-year return periods. Figures 6a–6e show footprints that resemble well known historic events, while Figures 6f–6j provide examples of other common footprints in which we see a distinct clustering of gauges (see figure caption for details). Figure 6 therefore suggests that the dependence model can produce event footprints for the United States that are at least intuitively realistic.

A more formal validation was undertaken by comparing the mean observed event footprint size at each gauge with the same quantity in the set of simulated events (Figure 7). Figure 7 plots the mean footprint size in the modeled and observed data sets at each gauge location. The mean is calculated from the footprints of all events that impacted the given gauge location with a flow magnitude greater or equal to the 99th percentile (i.e., Q99, the prediction threshold in this study). This analysis demonstrates that the model can reproduce the typical size of event footprints observed from the gauge record. It is worth noting that at least some of the differences between observed and modeled mean footprint size arise because the observed

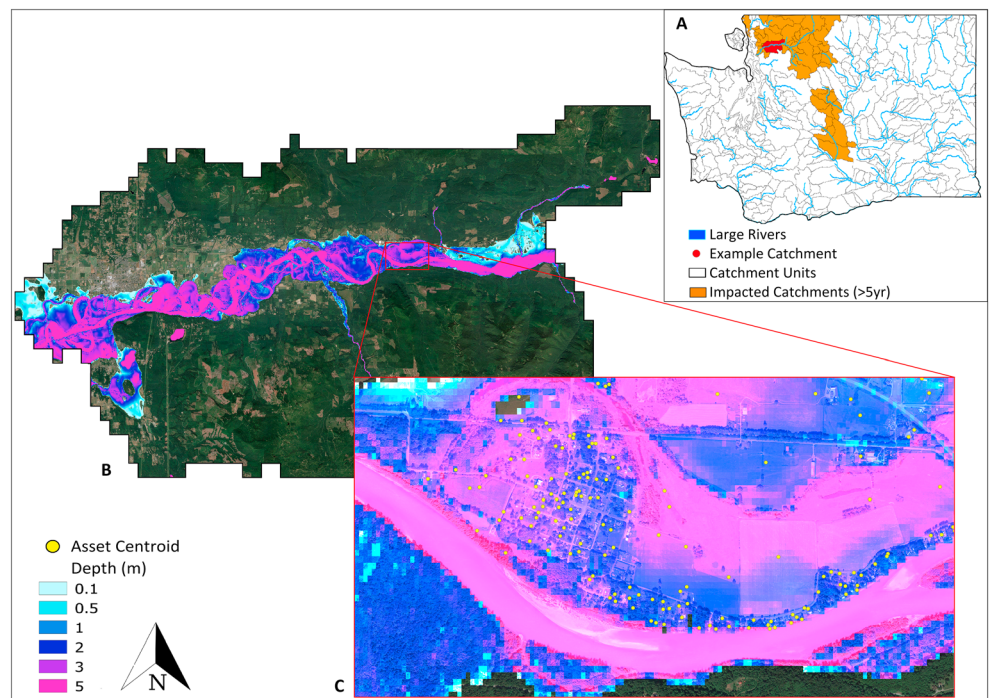


Figure 5. Converting event footprints into economic damage estimates. In panel (a) an event footprint is generated providing an estimated return level flow magnitude within rivers in each catchment. For simplicity, all catchments experiencing an event >1 in 5 years are shown in orange, while an example catchment to consider is shown in red. Panel (b) shows the example catchment experiencing a 100-year event with the expected inundation depths extracted from the corresponding continental hydrodynamic model hazard layer. The bottom panel (c) shows an example section of the impacted river reach with an overlay of the FEMA National Structure Inventory asset data set (yellow points representing individual building locations). The intersection of the assets and the hazard layer depths, in conjunction with relevant depth-damage functions from the U.S. Army Corps of Engineers, enables an estimate of direct economic damage to be made.

record is a much smaller and more incomplete sample of possible event footprints at each gauge than the synthetic record.

Next, we compared the observed return period flow for each gauge with return period flows determined from events in the synthetic record (see Figure 8). We restrict this analysis to the 10- to 50-year return period range as for this we have a very large sample of such events. Return period exceedance is calculated by extracting the annual maximum flows for each year within the observed and synthetic records and fitting GEV models to each set of data. These GEV relationships are then used to estimate the 1 in T year exceedance flows. This indicates that the statistical model is able to simulate synthetic event magnitudes that closely recreate the marginal distributions of the gauge records. The deviations between the modeled and observed flows are shown to increase with return period magnitude, as one might expect.

4. Results

The validation tests reported above demonstrate that the conditional exceedance model is able to simulate realistic magnitudes and spatial patterns of U.S. flooding. Stochastic sampling from this statistical model allows the creation of a much larger sample of U.S. flood footprints than is available in the observed record and therefore allows better sampling of the tails of the flood hazard distribution. Interrogating the patterns of spatial dependence in the stochastically generated event set can therefore yield insights into the nature of flood hazard and risk in the United States that cannot be obtained in any other way. In this section we therefore analyze what the conditional exceedance model tells us about patterns of flood spatial dependence and loss-exceedance distributions. In the analysis that follows we only consider events where flow for at least one

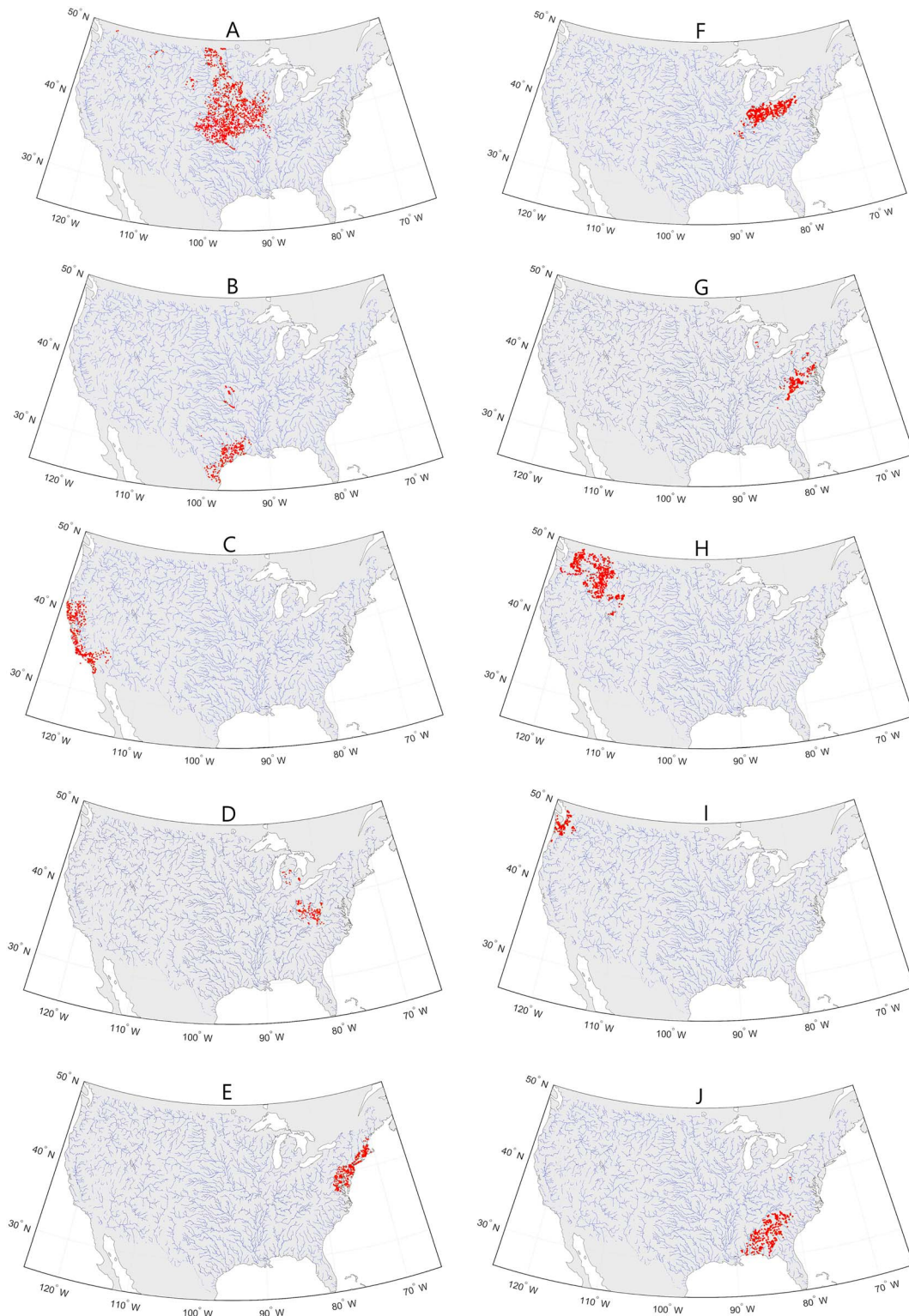


Figure 6. Example footprints found within the 1,000-year synthetic record. Red points refer to catchment centroids with magnitudes >1 in 5 years. For clarity, catchments with <5 -year return period flow during the example events have been excluded. Plots (a)–(e) show footprints that resemble well known historic events: Mississippi 1993 (a), Harvey 2017 (b), California floods of 2017 and 1997 (c), West Virginia 2016 (d), and Irene 2011 and Sandy 2012 (e). Plots (f)–(j) provide examples of other common footprints in which we see a distinct clustering of gauges: along the Ohio river, north of the Appalachian Mountains (f), clustering south of the Appalachian Mountains (g), large clusters of flooding in the Pacific North West (h), localized flooding in Western Washington (i), footprints running along a SW-NE transect along the east coast typical of flooding caused by tropical cyclones (j).

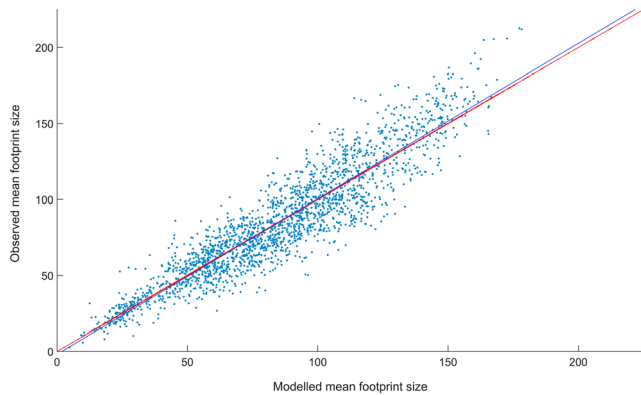


Figure 7. The mean footprint size in the modeled and observed data sets at each gauge point. Footprint size is given as the number of gauges experiencing a $\geq Q_{99}$ flow magnitude on the Laplace margin (the prediction threshold) when a given site of interest is also $\geq Q_{99}$. Means are calculated over all events at each gauge. The blue line indicates the linear least squares fit in the model while the red line represents the observed against observed (1:1) fit.

the footprint size (the area of catchments also experiencing the same event) also increases. Notable areas where this relationship is less strong are the Pacific North West coast, Florida, and coastal New England. Thus, while we would expect larger events to have larger footprints it is clear from the stochastic analysis that there is a geography to this relationship that is more significant than our current understanding of flood spatial dependence might have suggested.

Figure 10 looks at how the number of neighboring gauges also experiencing an event changes with increasing return period at a given site of interest. Here we plot for each gauge the mean number of neighboring gauges that experience flooding at >1 in 5-year return period when the return period at the conditioning site is >1 in 5 years, versus the equivalent numbers when the conditioning site experiences 5- to 20-, 50- to 100-, and >200 -year events. Like Figure 9, Figure 10 indicates that in nearly all cases the overall footprint size increases as the return period at the conditioning site increases. Additionally, Figure 10 shows that the variability in footprint size increases with return period, with the standard deviations from the group mean recorded as 1, 11, and 28 for the 5- to 20-, 50- to 100-, and >200 -year bins, respectively. This indicates that as conditioning site magnitude increases, so does the range in the expected footprint size. It should be noted that some of this variability may be associated with fewer events being recorded at higher magnitude return periods from which the means are calculated at each site.

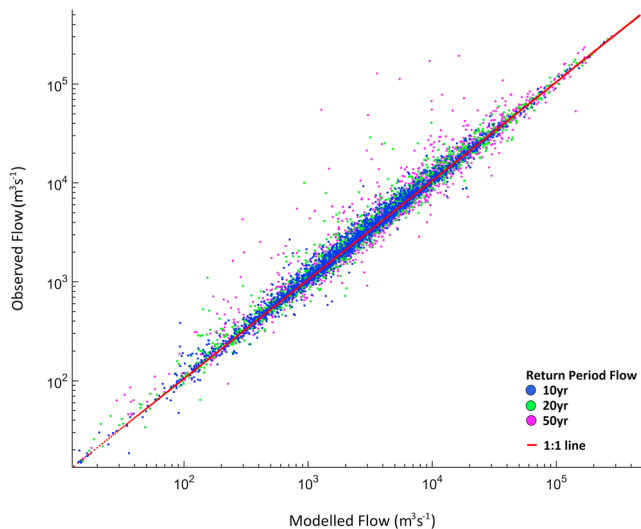


Figure 8. The 1 in 10- (blue), 20- (green), and 50-year (magenta) flow magnitudes estimated from the observed and synthetic time series data. In both instances, return period exceedance is calculated by extracting the annual maximum flows for each year within the records and fitting a Generalized Extreme Value model to the data. This is then used to estimate the 1 in T year exceedance flows. The red line indicates the 1:1 fit.

gauge site is greater than 1 in 5-year return period as floods smaller than this are unlikely to lead to significant losses.

4.1. Patterns of Flood Hazard

Figure 9 shows a map of the mean modeled footprint size in square kilometers for each gauge for all events in the synthetic record (Figure 9a) impacting a given site and for events within 5- to 20-, 50- to 100-, and >200 -year return period ranges, representing high, medium, and low frequency events (Figures 9b–9d, respectively). Each point in Figure 9 relates to a flow gauge location. Footprint size in this instance represents the mean area impacted during events that occur at the given site (for all events >5 -year return period, or within the specified bin size). The footprint size is the sum of all catchment areas (note that this is not the flooded area) that experience at least a 1 in 5-year event magnitude during a given event. Figure 9 makes clear the advantage of using a stochastically generated event set to increase the sample size as mapping footprint sizes for >200 -year return period events would be impossible from the observed record because of the limited sample size. The plots indicate that as the magnitude of an event at a given site increases, we typically find that

the footprint size (the area of catchments also experiencing the same event) also increases. Notable areas where this relationship is less strong are the Pacific North West coast, Florida, and coastal New England. Thus, while we would expect larger events to have larger footprints it is clear from the stochastic analysis that there is a geography to this relationship that is more significant than our current understanding of flood spatial dependence might have suggested.

Lastly for the hazard, Figure 11 looks at the distribution of event magnitudes within each footprint as the return period at the conditioning site changes. To examine this, we find the maximum flow magnitude in each synthetic event and then calculate the total event footprint size (number of gauges with ≥ 1 in 5-year flow), the number of gauges with a flow magnitude $\geq 50\%$ of the largest flow magnitude, and lastly the ratio of these two figures. This gives the proportion of the event footprint where flow is also relatively large. Figure 11 plots these results as histograms for four varying bins of maximum flow magnitude during a given event: (a) >1 in 5-, (b) 5- to 20-, (c) 50- to 100-, and (d) >200 -year. The plots indicate that, despite the total number of a gauge's neighbors that experience flooding increasing with increasing conditioning site flow magnitude, the proportion of neighboring gauges that experience return periods within 50% of the event maximum shows a decreasing trend. For instance, the mean number of gauges with flow $\geq 50\%$ of the conditioning site flow in the >200 -year bin was 1.7 despite the average footprint size being 59 gauges,

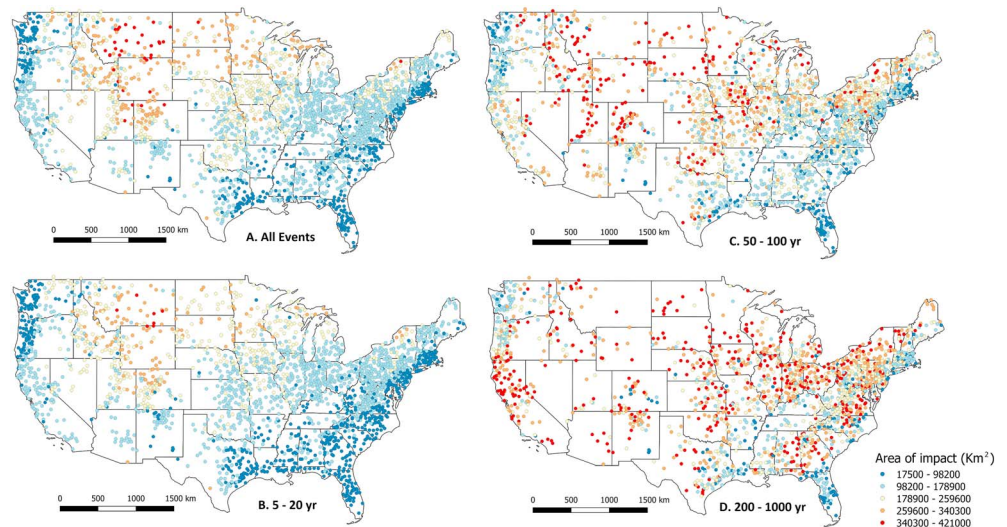


Figure 9. Mean modeled footprint area when a specified gauge experiences an event with return period within one of four different magnitude bounds (>5 , 5–20, 50–100, 200–1,000 years, shown in panels (a) to (d), respectively). The area is defined as the catchment area (not the area of inundation) of all catchments assigned a return period of greater than 5 years during a given event.

while the same assessment of the 5- to 20-year bin set gave a mean number of neighbors with flow $\geq 50\%$ of that at the conditioning site as 6.8 given an average footprint size of 14 gauges. Thus, as event magnitude becomes more extreme then equivalently extreme flows are experienced over an increasingly small proportion of the event footprint.

4.2. Patterns of Economic Loss

It is not just the hazard that is important to many end users of flood analyses, but rather it is the intersection of the hazard with assets and their corresponding vulnerabilities to obtain an economic damage estimate that is key. Typically, this is considered at an event or annual level, but validation of these values is notoriously difficult because most observed loss information is commercially privileged. One of the few exceptions to this is the annual loss estimates for inland flooding reported by NOAA's National Weather Service (NWS;

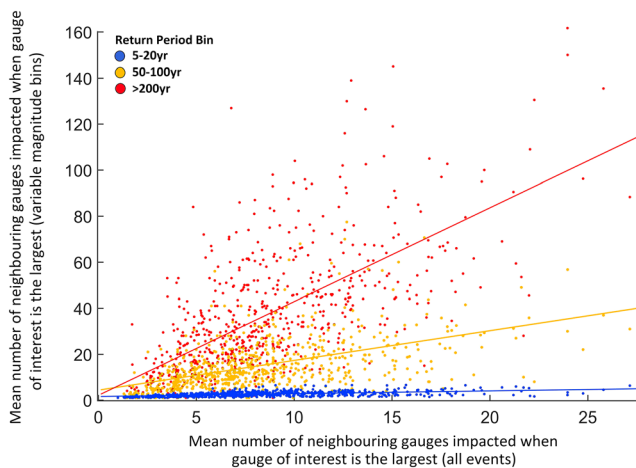


Figure 10. Scatterplots of the mean number of neighboring gauges impacted during an event at each gauge (x axis) in which the site of interest experiences an event with a flow magnitude >1 in 5 years, against the mean number of neighboring gauges impacted at the same gauge when the event maximum falls within one of three bins (y axis): 5–20 (blue), 50–100 (orange), >200 years (red). The plot indicates that as the event maximum flow return period increases, so too does the mean event footprint.

see <http://www.nws.noaa.gov/hic/>). These are compiled each year by NWS offices using data directly obtained from a variety of sources including emergency managers, the USGS, the U.S. Army Corps of Engineers, power utility companies, and newspaper articles or estimates indirectly based on established guidelines. The loss data for individual events are quality controlled and then compiled nationally to produce an annual total, with historic economic losses being adjusted each year for inflation. The NWS data report direct fluvial losses to the majority of infrastructure classes recorded in the U.S. National Structure Inventory that is used in the model loss calculations. The model similarly considers direct fluvial losses and therefore using the NWS data set for loss validation is as close to a like-for-like comparison as can be made using publicly available information. The NWS note that, in general, the historic data are less accurate than data obtained more recently, and that the estimates are subject to a wide variety of errors including underreporting and subjectivity. It is likely that underreporting will be a particular issue for smaller magnitude floods where those affected are either self-insured or may for other reasons not consider it sufficiently important or worthwhile to record the damage. Flood victims may also be concerned to not report loss information for smaller floods if they believe that this may affect the future value of their property or their future insurance premiums. Larger, more noteworthy events will likely be more carefully recorded. In an independent

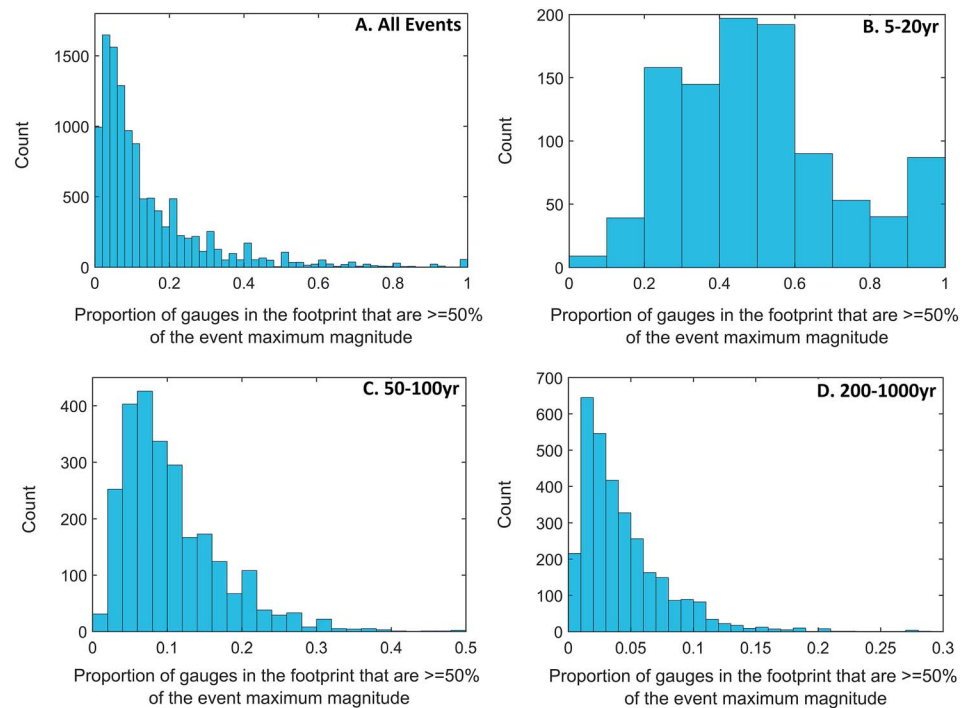


Figure 11. Histograms of the proportion of gauges within an event that are $\geq 50\%$ of the event maximum magnitude, given conditioning site magnitude falls within four bins, all events (a), 5–20 (b), 50–100 (c), and >200 years (d). The plots indicate that as event peak magnitude increases, the proportion of the event footprint that is relatively large (given here as the number of neighboring gauges experiencing flow $\geq 50\%$ of the maximum event flow) decreases.

analysis of the NWS data, Downton and Pielke (2005) found that losses for small events were often extremely inaccurate, but that independent estimates of losses for events with total damage greater than \$500 million disagreed by less than 8%. They concluded that estimates aggregated over large areas or over long time periods appeared to be reasonably reliable, thereby justifying the use of national annual losses in this paper.

Historic loss data will also necessarily misestimate the present-day loss-exceedance curve across all probabilities as a result of nonstationarity of climate and land use. Using observed losses from the past to evaluate simulated losses of the present assumes that historic floods occurred within the same societal context as today. This is an easily dismissed assumption, as rapid socioeconomic changes have occurred in recent U. S. history. According to the NWS record, flooding in 1993 caused over \$30Bn worth of damage, yet this took place when the Gross Domestic Product (GDP) of the United States was one third of what it is today (US Bureau of Economic Analysis, 2018). The synthetic losses generated by the model developed in this paper can be conceptualized as 1,000 simulations of possible flood damages in 2010, since this was when the National Structure Inventory was created. If our model was to perfectly replicate the flood events that occurred in 1993, it would produce a dramatically higher loss because of changes in exposure since then. To decouple historic losses from the societal conditions in which they took place, the NWS record must first be “normalized.” Using historic U.S. GDP from the US Bureau of Economic Analysis (2018) as a proxy for changes in exposure, as is common in normalization procedures (e.g., Barredo, 2009; Miller et al., 2008; Nordhaus, 2006; Pielke et al., 2008), we adjust the NWS figures to estimate the damage historical events would have caused under 2010 levels of development. For example, U.S. GDP in 2010 was 3.5 times U.S. GDP in 1985, so the inflation-adjusted NWS loss in 1985 of \$1.2Bn is normalized to \$4.2Bn. Neumayer and Barthel (2011) note that while temporal socioeconomic changes are accounted for in this approach, changes across space at a point in time are not. This is an unfortunate constraint of the data that are available, but across a long enough time period the assumption that historic disaster zones developed at similar rates to unaffected areas is likely to hold.

Figure 12 plots the probability of total annual flood damage exceedance across all U.S. assets obtained from the synthetic model record (red, GEV distribution). We analyze the NWS loss data in two different ways: the

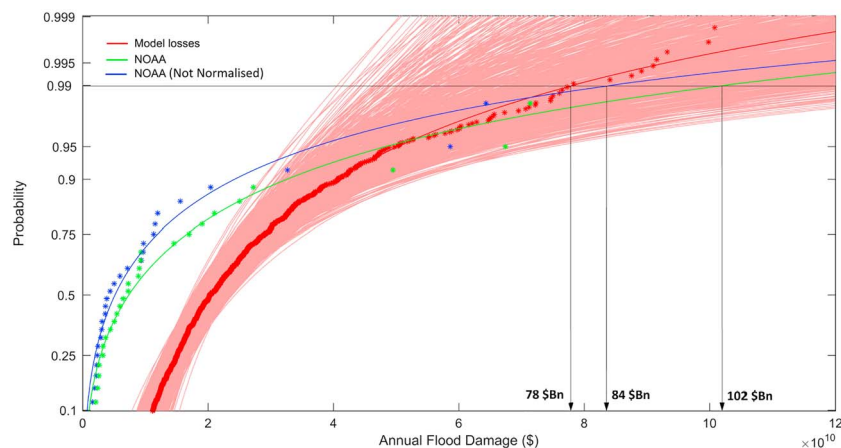


Figure 12. The normalized modeled probability of annual flood damage exceedance (red) plotted against two variants of the annual losses provided by NOAA National Weather Service: the 30-year period of data between 1988 and 2017 (blue), the same period but with values normalized based on Gross Domestic Product (green). Arrows indicate the estimated 1% exceedance annual loss from each distribution. The 1,000 samples of 30 years of loss were subsampled from the model ensemble and Generalized Extreme Value distributions fitted to each (each shown as a light red curve).

30-year period of data between 1988 and 2017 normalized for GDP growth (green) and the same 30-year period without normalization (blue). To each observed data set we fit a distribution, in this instance a Generalized Pareto distribution as this was best suited to the observed data. Despite leading to fewer data points with which to estimate a relationship, a focus on the last 30 years of the record is likely important as flood exposure and vulnerability will have changed significantly since the NWS records began. The plots indicate that, relative to the NWS estimates, the synthetic model has much larger loss values within the high frequency portion of the distribution (e.g., <1 in 10 year return period annual losses) but begins to converge within the range of estimates provided by the NOAA data at return periods above this. This is consistent with the details of the observed data capture methodology reported by NOAA and the above hypothesis that underreporting will be a more significant issue for smaller magnitude events. It is also likely that the hydraulic model overpredicts loss for high frequency events as these are more sensitive to the correct representation of flood defenses and channel capacity. Modeling small flood events is also inherently difficult with anything other than airborne LiDAR terrain data (e.g., data with 1–2 m resolution, ~10 cm RMSE vertical accuracy) because coarser DEMs, such as the 30-m information used here, do not fully capture the fine-scale floodplain flow pathways that control floodplain inundation and dewatering during smaller flood events (Bates et al., 2006; Neal et al., 2011; Ozdemir et al., 2013; Sampson et al., 2012). The 1% exceedance probability (often the focus in flood risk studies) is also shown in the figure. If we consider only the last 30 years of NWS data then the observed 1% annual probability loss for U.S. fluvial flooding is ~\$102Bn (GDP normalized, ~\$84Bn without normalization) compared to ~\$78Bn in the synthetic event set. Given probable errors in both the observed and modeled data this difference is unlikely to be statistically significant.

The ability of historic observations to accurately represent the tail of the loss-exceedance curve is also impaired by the length of the record used. The most recent 30 years of observations are typically used to ensure that only flood events of the present climate (defined by the World Meteorological Organization as a period of three decades) are considered, which therefore brings into question the ability of even normalized losses to capture extremely rare and damaging flood events (e.g., 1 in 100-year floods or greater). A distribution of 1,000 synthetic realizations of loss in 2010 will necessarily look different to one consisting of only 30 observed losses normalized to 2010 societal conditions, since the former has the freedom (in a crude temporal sense) to explore extremes and hence better define the tail. In light of this, further to comparing our 1,000-year loss-exceedance curve to that of the normalized losses, we also consider 1,000 random 30-year samples from our event set. These alternative realizations of the 30-year loss profile (GEV distribution) are shown in Figure 12 and indicate the uncertainty range in the modeled losses as a result of the random sampling of flood events over a short record length. It follows that a similar uncertainty range is very likely to hold for the observed data too as a result of this factor alone.

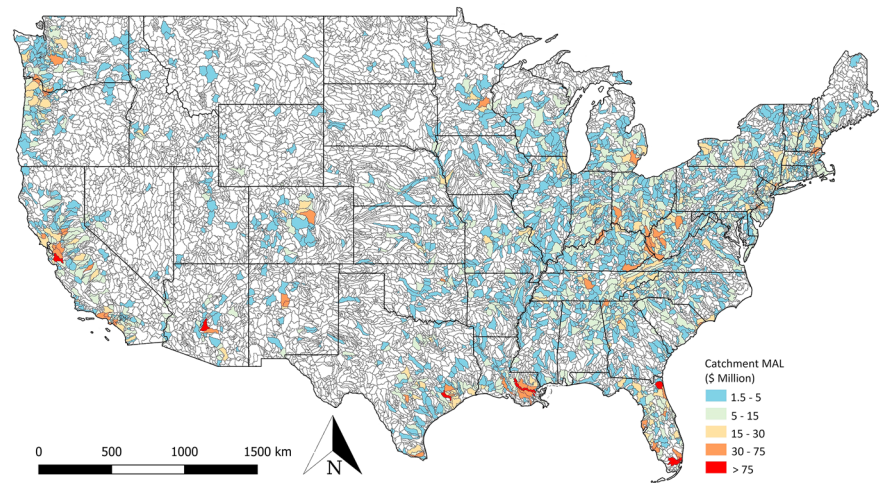


Figure 13. Modeled catchment mean annual loss. For clarity, only catchments with a mean annual loss of $> \$1.5$ million have been plotted.

Figure 13 considers the spatial variability in expected damages across the United States, plotting the modeled mean annual loss associated with each catchment as a fraction of the national total. In other words, Figure 13 shows the catchments which are hotspots of flood risk. The results indicate that the greatest contributions to mean annual national loss are associated with catchments in densely populated regions, with relatively high loss catchments found in areas such as Houston, Louisiana, Miami, San Francisco, and Seattle. This indicates that, despite the expectation of relatively high design standard flood defenses being present in these regions, they are still the zones of highest risk.

Figure 14 looks at broader patterns of risk and hazard concentration. Here we plot the number of times large loss causing events (defined here as those exceeding the 0.99 quantile event loss of $\$5.5$ Bn) impact a given catchment. This is subtly different to the spatial pattern of absolute risk in Figure 13 because a catchment may be part of large loss causing event but not contribute significantly to the total event loss. Figure 14 does, however, show the broad regions in which large loss causing events are more frequent. Like Figure 13 the results indicate that extreme losses are associated with events that impact catchments that contain densely populated centers such as California and Florida. However, Figure 14 also shows an extensive region of catchments from West Virginia through to New York that broadly follow the Appalachian mountain chain and which experience flooding during large loss causing events.

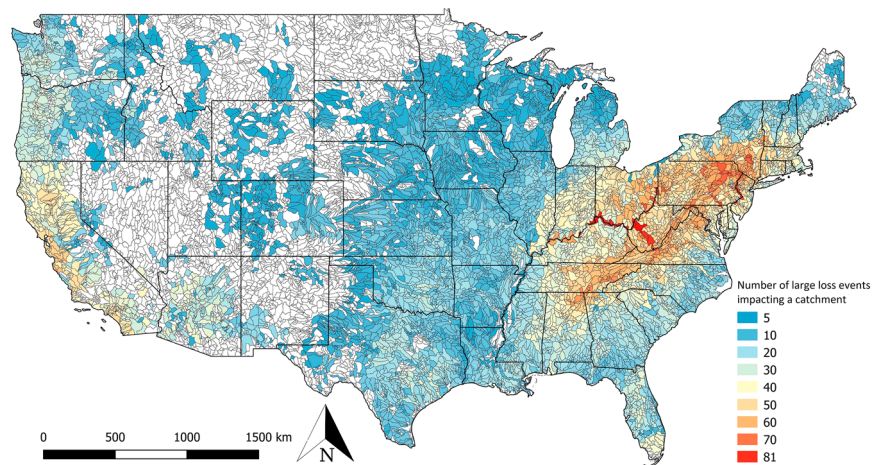


Figure 14. The number of large loss causing events impacting catchments. In this instance large loss causing events are those associated with damages $> \$5.5$ billion (the 0.99 probability event loss threshold).

5. Discussion

This analysis of spatial dependence during flood events over a large and heterogeneous region yields a number of insights that add to our current knowledge of U.S. flood hazard and risk. It also seems likely that similar effects would be observed in other territories as well. First, while it is reasonably intuitive that event footprint size should increase with event peak magnitude, there appears to be an unexpectedly strong geography to this (Figure 9). This would seem to hold despite the range in possible event footprint size also becoming much wider with increasing event peak return period (Figure 10). For example, Figure 9 shows that in the Pacific North West the expected pattern of increasing footprint size with event magnitude still holds, but that this relationship is markedly less pronounced than in adjacent coastal California. It should be noted that spatial variations in gauge density may also have an unavoidable impact here. For example, in many gauge-poor areas of the mid-west the footprint size during an event is large, yet we might expect numbers of floods here to be caused by local convective weather systems that have small spatial footprints. However, for the Pacific North West/Coastal California example noted above the gauge density is likely sufficient for this to be a real effect. Moreover, the impact of gauge density effects on national-scale risk estimates is likely to be low due to typically lower asset densities in gauge-poor regions.

Second, we observe that as events become more extreme the area over which extreme flows take place becomes an increasingly smaller proportion of the total footprint size (Figure 11), even though the footprint area may be bigger in absolute terms. This has also not previously been noted in either empirical or modeling studies. This suggests that very extreme events tend to experience the highest flow magnitudes over a small “core” area rather than being extreme everywhere. Further, while total footprint size may increase with event magnitude, the size of the extreme core does not expand at the same rate. This is a clear contrast with the constant in space return period assumption used in most large-scale risk analysis to date and suggests that a spatially constant return period becomes increasingly unjustifiable as maximum event magnitude increases.

Third, it is important to note that the simulated loss estimates were obtained without calibration or tuning of either the model or the vulnerability functions that are used to convert flow depth to damage. Instead, we use standard U.S. Army Corps of Engineers depth-damage relationships “off the shelf” for this and the U.S. National Structure Inventory to identify asset locations. By contrast, common practice in catastrophe modeling is to calibrate predicted losses to observations by adjusting the depth-damage functions used (Guin, 2017). The reasonable fit to the NWS loss data obtained without calibration, at least for >1 in 10-year return period national annual losses, gives some confidence that the maps of risk concentration that the approach generates are sensible. Of particular interest here is the analysis of catchments which experience flooding during large loss causing events (total event losses $> \$5.5\text{bn}$). This shows two large regional concentrations: one in coastal California, and a much larger one, along the line of the Appalachian Mountains (Figure 14). Figure 14 adds important information to our understanding of the spatial distribution of flood risk in the United States and also makes clear the linkages between terrain, orographic precipitation enhancement, and exposure.

Lastly, a motivation for switching from a “constant return period in space” approach to loss estimation from a stochastically generated event set was to enable a wider range of national-scale flood risk management questions to be properly addressed. For example, from a set of different magnitude constant return period national hazard layers it is not possible to compute a realistic value of the 1% annual probability national total loss. One can use a set of constant return period layers to calculate an expected AAL by assuming each return period layer contributes a pro rata proportion of the total AAL value, but only a set of stochastically generated flood footprints with correct spatial dependence can correctly determine the risk profile, that is, the loss-exceedance curves shown in Figure 12. For example, estimates for the loss that would result from a 1 in 100-year flood event happening across the whole United States simultaneously (the constant return period in space assumption) range from \$1.2 to \$3.6Tn (Wing et al., 2018; Winsemius et al., 2013). This event frequency should then, on average, contribute 1% of this value (i.e., $\sim \$12\text{--}36\text{Bn}$) to the AAL. By performing the same calculation for a range of event magnitudes an estimate of the overall AAL can be made, although this is likely to be significant overestimate given that the observed AAL in the NWS data is only $\sim \$14\text{Bn}$, normalized, over the last 30 years. An AAL analysis can only determine what happens in an “average” year and cannot assess the year-to-year variability in national annual loss that results from the stochastic nature of

flood event occurrence. It also cannot determine a realistic magnitude for different exceedance probability losses, e.g., the 1% U.S. annual flood loss of ~\$78Bn given above, yet numerous national-scale flood risk management decisions require exactly this kind of information.

While this study has yielded a number of new insights into the nature of flood spatial dependence and risk for the United States some caveats and areas for further work need to be noted. In particular, the definition used to determine what constitutes an independent event is necessarily somewhat arbitrary. It is clear that there will be some events that are too long or too widespread to be properly captured in the time and space window we have selected, while at the same time some of our single “floods” will be spurious mixtures of multiple events. While different methodological choices could have been made (and perhaps justified), the same general problem will always remain. A similar unavoidable issue is the impact that spatial variations in gauge density will have on the results, although the typically positive correlation between gauge and asset density does mitigate this effect to some extent. Lastly, in this first version of a conditional multivariate exceedance model for the United States we do yet account for the persistence of unusual meteorological or hydrological conditions, which means it does not properly represent event clustering. Neither do we represent the effects of seasonality on mixtures of storms (as in Schneeberger et al., 2018). By mixing all types of event together in the multivariate exceedance model we are likely to be confusing this signal, especially in areas that experience a mixture of high flows due to different types of flood causing events at different times of year. It will be useful and important to address both issues in later versions of the framework.

6. Conclusions

This paper has presented a first comprehensive analysis of spatial dependence in U.S. flood hazard and risk using an established conditional multivariate exceedance statistical model (Heffernan & Tawn, 2004) applied to ~2,400 USGS river gauge sites to produce a stochastic catalogue of >63,000 flood events with realistic spatial dependence. Previously, such methods have only been applied to small and relatively homogeneous regions and extensions to the method were required to deal with this additional complexity and also the increased computational cost. A new hydrodynamic model of the conterminous United States at 30-m resolution (Wing et al., 2017) was used to calculate the loss arising from each of these synthetic events and thereby derive a national-scale loss-exceedance curve for fluvial flooding in the United States. This curve was shown to be a reasonable match to an observed loss-exceedance curve derived for the period 1984–2014 from data provided by the U.S. NWS, at least for annual loss-exceedance probabilities smaller than 10% (or 1 in 10-year return period). Importantly, this result was achieved without calibration of the vulnerability functions using observed losses.

Analysis of the set of stochastic events and losses yields new insights to the nature of flooding and flood risk in the United States. While event footprint size (sum of all catchments areas affected by flooding) increases with increasing event peak magnitude, there appears to be an unexpectedly strong geography to this: Adjacent well-gauged regions show very different relationships between event peak magnitude and footprint size. Events with larger peak magnitude show greater variability in footprint size and the area over which extreme flows take place becomes an increasingly smaller portion of the total footprint size as event peak magnitude increases (even if the absolute footprint size is increasing). We find that the greatest contribution to mean annual national loss is associated with catchments in densely populated regions, despite the expectation that these are protected by relatively high design standard flood defenses. Lastly, analysis of catchments impacted during single large loss causing events (losses >\$5.5Bn) shows concentrations in coastal California, Florida, and along the line of the Appalachian Mountains. This highlights the linkages between terrain, orographic precipitation enhancement, exposure, and risk. While further work is required to include both seasonality effects and storm clustering into the framework, the analysis provides a significant advance over previous national-scale risk analyses that assume flood return periods that are constant in space.

Acknowledgments

Paul Bates was supported by a Research Fellowship from the Leverhulme Trust and a Royal Society Wolfson Research Merit Award. Jeff Neal was supported by UK Natural Environment Research Council grant NE/S006079/1. The USGS gauge data used in the paper are available online (<https://waterdata.usgs.gov/nwis>). The U.S. National Elevation Data Set used in the hydrodynamic model are available online (<http://www.ned.usgs.gov>). The event set and hydrodynamic model data layers are proprietary in nature but can be made available for research use (i.e., not commercial, policy, or regulatory applications) by contacting Fathom Ltd. (info@fathom.global).

References

- Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., & Feyen, L. (2013). Advances in pan-European flood hazard mapping. *Hydrological Processes*, 28(13), 4067–4077. <https://doi.org/10.1002/hyp.9947>
- Allen, G. H., David, C. H., Andreadis, K. M., Hossain, F., & Famiglietti, J. S. (2018). Global estimates of river flow wave travel times and implications for low-latency satellite data. *Geophysical Research Letters*, 45, 7551–7560. <https://doi.org/10.1029/2018GL077914>

- Almeida, G. A. M., & Bates, P. (2013). Applicability of the local inertial approximation of the shallow water equations to flood modeling. *Water Resources Research*, 49, 4833–4844. <https://doi.org/10.1002/wrcr.20366>
- Almeida, G. A. M., Bates, P., Freer, J. E., & Souvignet, M. (2012). Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. *Water Resources Research*, 48, W05528. <https://doi.org/10.1029/2011WR011570>
- Barredo, J. I. (2009). Normalised flood losses in Europe: 1970–2006. *Natural Hazards and Earth System Sciences*, 9(1), 97–104. <https://doi.org/10.5194/nhess-9-97-2009>
- Bates, P. D., & De Roo, A. P. J. (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1–2), 54–77. [https://doi.org/10.1016/S0022-1694\(00\)00278-X](https://doi.org/10.1016/S0022-1694(00)00278-X)
- Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, 387(1–2), 33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>
- Bates, P. D., Wilson, M. D., Horritt, M. S., Mason, D. C., Holden, N., & Currie, A. (2006). Reach scale floodplain inundation dynamics observed using airborne synthetic aperture radar imagery: Data analysis and modelling. *Journal of Hydrology*, 328(1–2), 306–318. <https://doi.org/10.1016/j.jhydrol.2005.12.028>
- Blöschl, G., & Sivapalan, M. (1997). Process controls on regional flood frequency: Coefficient of variation and basin scale. *Water Resources Research*, 33, 2967–2980. <https://doi.org/10.1029/97WR00568>
- Cabinet Office. (2017). UK National Risk Register 2017 (P. 71). London, UK. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/644968/UK_National_Risk_Register_2017.pdf
- Chow, V. T. (1959). *Open-channel hydraulics*. New York: McGraw-Hill.
- Cooley, D., Nychka, D., & Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479), 824–840. <https://doi.org/10.1198/016214506000000780>
- de Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E., & Ward, P. J. (2015). Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change*, 20(6), 865–890. <https://doi.org/10.1007/s11027-015-9654-z>
- Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., & Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. *Advances in Water Resources*, 94, 87–102. <https://doi.org/10.1016/j.advwatres.2016.05.002>
- Downton, M. W., & Pielke, R. A. (2005). How accurate are disaster loss data? The case of U.S. flood damage. *Natural Hazards*, 35(2), 211–228. <https://doi.org/10.1007/s11069-004-4808-4>
- Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hündecha, Y., et al. (2015). Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. *Journal of Hydrology*, 524, 182–193. <https://doi.org/10.1016/j.jhydrol.2015.02.021>
- Guin, J. (2017). What makes a catastrophe model robust? In G. Michel (Ed.), *Risk modeling for hazards and disasters* (pp. 47–62). The Netherlands: Elsevier.
- Gupta, V. K., Castro, S. L., & Over, T. M. (1996). On scaling exponents of spatial peak flows from rainfall and river network geometry. *Journal of Hydrology*, 187(1–2), 81–104. [https://doi.org/10.1016/S0022-1694\(96\)03088-0](https://doi.org/10.1016/S0022-1694(96)03088-0)
- Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 66(3), 497–546. <https://doi.org/10.1111/j.1467-9868.2004.02050.x>
- Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., et al. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States; representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, 81(5), 345–254.
- Horritt, M. S., & Bates, P. D. (2001a). Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, 253(1–4), 239–249. [https://doi.org/10.1016/S0022-1694\(01\)00490-5](https://doi.org/10.1016/S0022-1694(01)00490-5)
- Horritt, M. S., & Bates, P. D. (2001b). Predicting floodplain inundation: Raster-based modelling versus the finite-element approach. *Hydrological Processes*, 15(5), 825–842. <https://doi.org/10.1002/hyp.188>
- Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., et al. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change*, 4(4), 264–268. <https://doi.org/10.1038/nclimate2124>
- Keef, C., Svensson, C., & Tawn, J. A. (2009). Spatial dependence in extreme river flows and precipitation for Great Britain. *Journal of Hydrology*, 378(3–4), 240–252. <https://doi.org/10.1016/j.jhydrol.2009.09.026>
- Keef, C., Tawn, J. A., & Lamb, R. (2012). Estimating the probability of widespread flood events. *Environmetrics*, 24(1), 13–21. <https://doi.org/10.1002/env.2190>
- Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., et al. (2010). A new method to assess the risk of local and widespread flooding on rivers and coasts. *Journal of Flood Risk Management*, 3(4), 323–336. <https://doi.org/10.1111/j.1753-318X.2010.01081.x>
- Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, 27(15), 2171–2186. <https://doi.org/10.1002/hyp.9740>
- Lu, C., Singh, V. P., Guo, S., Zenchao, H., & Tianyuan, L. (2012). Flood coincidence risk analysis using multivariate copula functions. *Journal of Hydrologic Engineering*, 17(6), 742–755. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000504](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000504)
- Lu, P., Smith, J. A., & Lin, N. (2017). Spatial characterization of flood magnitudes over the drainage network of the Delaware River basin. *Journal of Hydrometeorology*, 18(4), 957–976. <https://doi.org/10.1175/JHM-D-16-0071.1>
- Miller, S., Muir-Wood, R., & Boissonade, A. (2008). An exploration of trends in normalized weather-related catastrophe losses. In H. F. Diaz, & R. J. Murnane (Eds.), *Climate extremes and society* (pp. 225–247). New York: Cambridge University Press. Retrieved from <https://www.cambridge.org/core/books/climate-extremes-and-society/an-exploration-of-trends-in-normalized-weather-related-catastrophe-losses/824B8A29C22EC8D3D2C7D8AF4A57FE1B>. <https://doi.org/10.1017/CBO9780511535840.015>
- Neal, J., Keef, C., Bates, P., Beven, K., & Leedal, D. (2012). Probabilistic flood risk mapping including spatial dependence. *Hydrological Processes*, 27(9), 1349–1363. <https://doi.org/10.1002/hyp.9572>
- Neal, J., Schumann, G., & Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, 48, W11506. <https://doi.org/10.1029/2012WR012514>
- Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P., & Mason, D. (2011). Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. *Journal of Flood Risk Management*, 4(2), 88–95. <https://doi.org/10.1111/j.1753-318X.2011.01093.x>
- Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). New York: Springer-Verlag. Retrieved from <http://www.springer.com/gb/book/9780387286594>
- Neumayer, E., & Barthel, F. (2011). Normalizing economic loss from natural disasters: A global analysis. *Global Environmental Change*, 21(1), 13–24. <https://doi.org/10.1016/j.gloenvcha.2010.10.004>

- Nordhaus, W. (2006). The economics of hurricanes in the United States (no. w12813). Cambridge, MA: National Bureau of Economic Research. doi: <https://doi.org/10.3386/w12813>
- O'Connor, J. E., & Costa, J. E. (2004). Spatial distribution of the largest rainfall-runoff floods from basins between 2.6 and 26,000 km² in the United States and Puerto Rico. *Water Resources Research*, 40, W01107. <https://doi.org/10.1029/2003WR002247>
- Ozdemir, H., Sampson, C. C., de Almeida, G. A. M., & Bates, P. D. (2013). Evaluating scale and roughness effects in urban flood modelling using terrestrial LIDAR data. *Hydrology and Earth System Sciences*, 17(10), 4015–4030. <https://doi.org/10.5194/hess-17-4015-2013>
- Pappenberger, F., Dutra, E., Wetterhall, F., & Cloke, H. L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrology and Earth System Sciences*, 16(11), 4143–4156. <https://doi.org/10.5194/hess-16-4143-2012>
- Pielke, R. A., Gratz, J., Landsea, C. W., Collins, D., Saunders, M. A., & Musulin, R. (2008). Normalized hurricane damage in the United States: 1900–2005. *Natural Hazards Review*, 9(1), 29–42. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2008\)9:1\(29\)](https://doi.org/10.1061/(ASCE)1527-6988(2008)9:1(29))
- Rougier, J. C. (2013). Quantifying natural hazard risk. In J. C. Rougier, R. S. J. Sparks, & L. J. Hill (Eds.), *Risk and uncertainty assessment for natural hazards* (pp. 19–39). UK: Cambridge University Press. <https://doi.org/10.1017/CBO9781139047562.003>
- Sampson, C. C., Fewtrell, T., Duncan, A., Shaad, K., Horritt, M. S., & Bates, P. D. (2012). Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in Water Resources*, 41, 1–17. <https://doi.org/10.1016/j.advwatres.2012.02.010>
- Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., & Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resources Research*, 51, 7358–7381. <https://doi.org/10.1002/2015WR016954>
- Schneeberger, K., Steinberger, T., Schneeberger, K., & Steinberger, T. (2018). Generation of spatially heterogeneous flood events in an alpine region—Adaptation and application of a multivariate modelling procedure. *Hydrology*, 5(1), 5. <https://doi.org/10.3390/hydrology5010005>
- Schumann, G. J. -P., Stampoulis, D., Smith, A. M., Sampson, C. C., Andreadis, K. M., Neal, J. C., & Bates, P. D. (2016). Rethinking flood hazard at the global scale. *Geophysical Research Letters*, 43, 10,249–10,256. <https://doi.org/10.1002/2016GL070260>
- Serinaldi, F., & Kilsby, C. G. (2017). A blueprint for full collective flood risk estimation: Demonstration for European River flooding. *Risk Analysis*, 37(10), 1958–1976. <https://doi.org/10.1111/risa.12747>
- Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11(12), 1649–1669. [https://doi.org/10.1002/\(SICI\)1099-1085\(19971015\)11:12<1649::AID-HYP495>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1099-1085(19971015)11:12<1649::AID-HYP495>3.0.CO;2-1)
- Smith, A., Sampson, C., & Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, 51, 539–553. <https://doi.org/10.1002/2014WR015814>
- Smith, J. A., Bradley, A. A., & Baeck, M. L. (1994). The space-time structure of extreme storm rainfall in the southern plains. *Journal of Applied Meteorology*, 33(12), 1402–1417.
- Smith, J. A., Villarini, G., & Baeck, M. L. (2011). Mixture distributions and the hydroclimatology of extreme rainfall and flooding in the eastern United States. *Journal of Hydrometeorology*, 12(2), 294–309. <https://doi.org/10.1175/2010JHM1242.1>
- Speight, L. J., Hall, J. W., & Kilsby, C. G. (2017). A multi-scale framework for flood risk analysis at spatially distributed locations. *Journal of Flood Risk Management*, 10(1), 124–137. <https://doi.org/10.1111/jfr3.12175>
- Taleb, N. N. (2007). *The Black Swan: The impact of the highly improbable*. UK: Random House. Retrieved from <https://www.penguin.co.uk/books/56380/the-black-swan/>
- Timonina, A., Hochrainer-Stigler, S., Pflug, G., Jongman, B., & Rojas, R. (2015). Structured coupling of probability loss distributions: Assessing joint flood risk in multiple river basins. *Risk Analysis*, 35(11), 2102–2119. <https://doi.org/10.1111/risa.12382>
- Turner-Gillespie, D. F., Smith, J. A., & Bates, P. D. (2003). Attenuating reaches and the regional flood response of an urbanizing drainage basin. *Advances in Water Resources*, 26(6), 673–684. [https://doi.org/10.1016/S0309-1708\(03\)00017-4](https://doi.org/10.1016/S0309-1708(03)00017-4)
- UNISDR (2015). *Making development sustainable: The future of disaster risk management*. Geneva: United Nations.
- US Bureau of Economic Analysis. (2018). Gross domestic product. Retrieved August 29, 2018, from <https://fred.stlouisfed.org/series/GDP/>
- Villarini, G., & Smith, J. A. (2010). Flood peak distributions for the eastern United States. *Water Resources Research*, 46, W06504. <https://doi.org/10.1029/2009WR008395>
- Villarini, G., Smith, J. A., Baeck, M. L., Marchok, T., & Vecchi, G. A. (2011). Characterization of rainfall distribution and flooding associated with U.S. landfalling tropical cyclones: Analyses of Hurricanes Frances, Ivan, and Jeanne (2004). *Journal of Geophysical Research*, 116, D23116. <https://doi.org/10.1029/2011JD016175>
- Vogel, R. M., Zafirakou-Koulouris, A., & Matalas, N. C. (2001). Frequency of record-breaking floods in the United States. *Water Resources Research*, 37, 1723–1731. <https://doi.org/10.1029/2001WR900019>
- Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research*, 53, 7968–7986. <https://doi.org/10.1002/2017WR020917>
- Wing, O. E. J., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione, J., & Morefield, P. (2018). Estimates of present and future flood risk in the conterminous United States. *Environmental Research Letters*, 13(3), 034023. <https://doi.org/10.1088/1748-9326/aaac65>
- Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., & Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, 17(5), 1871–1892. <https://doi.org/10.5194/hess-17-1871-2013>
- Wyncoll, D., & Gouldby, B. (2013). Integrating a multivariate extreme value method within a system flood risk analysis model. *Journal of Flood Risk Management*, 8(2), 145–160. <https://doi.org/10.1111/jfr3.12069>
- Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, 47, W04501. <https://doi.org/10.1029/2010WR009726>